DOCUMENT RESUME

ED 390 246 EC 304 511

AUTHOR de la Cruz, Rey E.
TITLE Assessment-Bias Issues in Special Education: A Review of Literature.
PUB DATE [96]
NOTE 37p.
PUB TYPE Information Analyses (070)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Ability Identification; *Court Litigation; Culture Fair Tests; *Disabilities; *Disability Identification; Elementary Secondary Education; Intelligence Quotient; *Intelligence Tests; Labeling (of Persons); Student Evaluation; *Student Placement; *Test Bias; Testing Problems; Test Validity

ABSTRACT

This paper reviews the literature on assessment bias issues in special education. While assessment instruments yielding a single IQ score are seen as useful components in a comprehensive multifactored assessment, and are the primary tool of diagnosis for mental retardation, they are found to be irrelevant when applied to students with learning disabilities (LD), as many students with LD have specific deficits in the areas most emphasized by IQ tests. Recent litigation in four cases of alleged assessment bias is reviewed, highlighting the history of controversy over racial differences in IQ and disproportionately high placements of students from racial minorities in special education, as well as low levels of minority placement in programs for gifted and talented students. The validity of using IQ tests for placement purposes arises as a major issue in these cases, as does the stigma associated with labelling. Principles of culture-balanced testing are explored, and while arguments in favor of completely eliminating IQ testing in schools are presented, surveys are reported to find broad-based support for IQ testing among psychologists and educational specialists. Alternatives to IQ testing are discussed, with an emphasis on multiple intelligences models and dynamic assessment. (Contains 67 references.) (PB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 390 246

ASSESSMENT-BIAS ISSUES IN SPECIAL EDUCATION: A REVIEW OF LITERATURE

REY E. DE LA CRUZ

EC 304511

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. Cruz

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

BEST COPY AVAILABLE

ASSESSMENT-BIAS ISSUES IN SPECIAL EDUCATION: A REVIEW OF LITERATURE

The term assessment was defined by Salvia and Ysseldyke (1995) as "the process of collecting data for the purposes of making decisions about students" (p. 5). Test data may be some of the information collected. McLoughlin and Lewis (1994) had a more specific definition: "The systematic process of gathering educationally relevant information in order to make legal and instructional decisions about the provision of special services to students with disabilities" (p. 601).

Assessment is big business in American schools. More than 20 million standardized tests are annually conducted on 44 million students in elementary and secondary schools (Ysseldyke, Algozzine, & Thurlow, 1992). The assessment market in special education is large. Students with disabilities are diagnosed, categorized, and instructed based on assessment results. In the school year 1991-1992, there were 4.9 million people under 22 years old who were in special education (U.S. Department of Commerce, 1994).

The use of IQ in determining students' eligibility for special education is a big concern. What is open to question is the appropriateness of IQ testing to students'

culture. The large number of minority students in special education programs is a problem (Maheady, Towne, Algozzine, Mercer, & Ysseldyke, 1990; Ysseldyke et al., 1992).

Assessment bias in special education is part of the larger debates about race, intelligence, and inequality in society (Herrnstein & Murray, 1994; Mensh & Mensh, 1991; Snyderman & Rothman, 1988). With the U.S. population becoming more diverse and multiculturalism being promoted, it is expected that assessment bias in special education will become more controversial. Caught in the middle of controversy are students with disabilities, whose needs ideally should be served by the assessment process.

Given the above concerns, there is a need to present the current research on how the problem is being addressed. The purpose of the study is to review the literature on assessment-bias issues in special education.

Identification and Evaluation

Public Law 94-142, the Education for All Handicapped Children Act of 1975, mandated free and appropriate public education for children, ranging in age between 3 and 21, who have disabilities (Mercer, 1991). Children with disabilities were defined as those with mental retardation, hardness of hearing, deafness, speech impairment, visual disability, serious emotional disturbance, orthopedic

impairment, other health impairments, deafness-blindness, multiple disabilities, or specific learning disabilities. In 1990, the Individuals With Disabilities Education Act (IDEA) included autism and traumatic brain injury in the definition of disabilities (Podemski, Marsh, Smith, & Price, 1995).

Testing, classification, and placement in special education programs are unavoidable. Children with disabilities are identified and evaluated by schools for the following reasons:

> (1) Plan, program, and appropriate funds for them, (2) provide appropriate services for them, (3) comply with federal and state laws requiring that disabled children be counted and served, (4) evaluate their own efforts to educate children. and (5) serve their bureaucratic interests in maintaining and expanding their own services. (Turnbull, 1993, p. 84)

The U.S. Congress required nondiscriminatory evaluation procedures. Consequently, instruments and procedures are used at a minimum by professionals to ensure nondiscriminatory assessment. First, tests and other evaluation materials should be conducted in the child's native language or other appropriate mode of communication. Second, test and evaluation materials should show areas of academic needs rather than just IQ. Third, test should demonstrate the child's aptitude and not the child's deficits. Fourth, different procedures have to be used to

determine an appropriate program. Fifth, a multidisciplinary team should participate in the assessment. Sixth, the child should be assessed in all aspects of the disability in question (Podemski et al., 1995; Turnbull, 1993).

Despite criticisms, IQ tests continue to be used in the classification of students with disabilities. Ideally, an intellectual assessment should only be a part of a multifactored assessment; other important information has to be considered. Intellectual assessment should play a role but not dominate the comprehensive evaluation (Reschly & Grimes, 1990). However, the inclusion of intelligence in the definitions of learning disabilities and mental retardation makes IQ tests indispensable.

The levels of IQ carry great weight in the classification of individuals with mental retardation. Beirne-Smith, Patton, and Ittenbach (1994) quoted the American Association of Mental Deficiency (AAMD) 1992 definition of mental retardation as "significantly subaverage intellectual functioning, existing concurrently with related limitations in two or more . . . applicable adaptive skill areas" (p. 75). Significantly subaverage means an IQ of 70 or below. Intended as a guideline, the upper limit can be extended to 75 or more, depending upon

the reliability of the IQ test used (Grossman, 1983). The flexibility of the ceiling score gives allowance for potential measurement error present in the specific test used (Polloway & Smith, 1988). Also, adaptive behaviors have to be considered. A student with an IQ score higher than 70 may be classified to have mental retardation because of deficient adaptive behaviors (Hunt & Marshall, 1994).

The primary criterion in identifying students with learning disabilities is the discrepancy between achievement and intelligence (Mercer, 1991). However, intelligence, as measured by IQ tests, was found by Siegel (1989) to be irrelevant to the definition of learning disabilities. IQ tests are inaccurate. They measure factual knowledge, expressive language abilities, and short-term memory, together with other skills. Because children with learning disabilities basically have deficits in these three areas, their low scores may be false. Also, low IQ scores are equated with poor reading. However, some children with low IQ scores have good reading scores. Thus, children with low IQ scores who are unsuccessful in reading have a reading disability. Their failure to read cannot be attributed to low IQ scores.

Placement Litigation

Court decisions on assessment bias have influenced the direction of special education. Nowhere is the issue more contested than in placement litigation.

In Diana v. State Board of Education (1970), the California Department of Education agreed to do the following: (a) to test bilingual children in both English and their primary language; (b) to delete unfair verbal items from tests; (c) to reevaluate all Mexican-American and Chinese students enrolled in classes for individuals with educable mental retardation, using nonverbal items and testing them in their native language; and (d) to make IQ tests that incorporate Mexican-American culture and are standardized only on the Mexican-American population (Salvia & Ysseldyke, 1995).

Lora v. Board of Education of the City of New York (1977, 1980, 1984) centered on the disproportionate number of African-American and Hispanic-American students in classes for individuals with emotional disturbance. The due process rights related to linguistic, cultural, or ethnic background differences of students were incorporated into the standards and procedures for nondiscriminatory assessment and decision making. A special advisory panel of experts on special educational programming was appointed.

Because decisions about the existence of emotional/behavioral disorders were subjectively made, professional responsibility was emphasized (Wood, Johnson, & Jenkins, 1986).

The Larry P. v. Riles (1972, 1979, 1984) case, which began in 1971, brought attention to assessment bias. The plaintiffs, who were six African-American children in the San Francisco Unified School District (SFUSD), complained about the unconstitutional standardized intelligence tests that placed a disproportionate number of African-American students in classes for individuals with educable mental retardation. For instance, from 1968 to 1969, African-American children comprised 27% of the population with educable mental retardation. It was a large number considering that only 9% of the California population was composed of African-American children. The city and state superintendents, the members of the State Board of Public Instruction, and the members of the City Board of Education were defendants in the case (Underwood & Mead, 1995). As a result, the state of California was ordered to stop using any standardized intelligence test for the identification of African-American students with educable mental retardation unless it was proven to be free of racial or cultural bias. In addition, all tests were to be conducted in a

nondiscriminatory way. Comparative African-American and European-American data on referrals and placement decisions were to be maintained. Disproportionate placements were to be eliminated and monitored. All African-American students were to be reevaluated without using standardized intelligence/ability tests (McLouglin & Lewis, 1994).

Larry P. was a premier case because it not only involved assessment bias, but also implied the stigma associated with classification or labelling, which may cause educational failure. The relationships among state legislation, state department of education policy, local school system implementation of policy, and the practices, procedures, and methods of special education and school psychology were demonstrated (Prasse & Reschly, 1986). Courts were used to answer scientific questions and settle matters of educational inequality (Prasse, 1988).

In analyzing Larry P. and other cases, Reschly, Kicklighter, and McKee (1988) corrected the misconception of predictive validity as the main criterion for individual IQ tests:

> We are absolutely certain that judicial scrutiny of IQ test use will be more positive if we communicate accurately about when IQ tests are used, that is, after severe, chronic low achievement leading to referral rather than prior to ample opportunities within the regular curriculum, and how IQ test results are used, that is, to confirm or disconfirm a hypothesis

> about low academic aptitude rather than a decision on whether or not the student should be exposed to the regular education curriculum. (p. 44)

Students typically spend a minimum of 2 years in general education before they are referred, and then, if appropriate, classified and placed in programs for individuals with educable mental retardation. IQ tests are therefore not used in a predictive validity sense, that is, "predicting" even before students are exposed to the general education curriculum.

MacMillan and Balow (1991) focused on incongruities in the state of California. First, the criteria for student with mental retardation and learning disabilities require "subaverage general intellectual functioning" and severe discrepancy between aptitude and achievement, respectively. Yet the scales that measure "intellectual functioning" or "attitude" cannot be used. Second, African-American students cannot be given a test of intelligence for Gifted and Talented Education (GATE) services; standardized test of achievement may be used. However, African-American students who have high aptitude but lacking "talents" may be excluded from GATE services. Third, aptitude and achievement tests tend to correlate; therefore, differences in school achievement should be emphasized rather than bias in measurement. Fourth, the disproportionate number of

African-American students in classes for individuals with educable mental retardation cannot be concluded in the context of total enrollment. For example, the disproportionate number of males is larger than the disproportion of African-American students. Fifth, the Larry P. decision applies only to African-American children, but many other children have biracial backgrounds.

MacMillan, Hendrick, and Watkins (1988) found that the interests of minority children were not well-served with the favorable rulings on the Diana and Larry P. cases. National statistics showed the continued overrepresentation of African-American children in classes for individuals with mild mental retardation. Special Education Pupil Count data for California, 1981-1982 (Incidence of Racial Isolation, 1983), revealed that 17.5% African-American children were classified to be with mental retardation, though they comprised only 9 to 10% of the school population. Many minority students attended segregated and inferior schools. They were further set back by their competency tests. Because of identification changes, the number of children in classes for individuals with mild mental retardation had decreased. The AAMD defined the IQ range to be from 50-55 to approximately 70. As such, children with IQs of 70-85, who were classified to be with mild mental retardation in

the past, were mainstreamed, i.e. placed in an an environment with their peers who have no disabilities (Overton, 1992). They were also served in other programs, e.g., learning disabilities. However, previous studies noted that a large number of children could not meet the criteria for learning disabilities. In California, the required intelligence test scores for severe discrepancy in learning disabilities were too high. Consequently, many minority children were classified as low achievers and never received special education services (MacMillan et al., 1988).

Larry P. was revisited in 1993. Parents of African-American students complained that the prohibition of using intelligence tests was discriminatory because African-American students did not meet learning disabilities criteria, one of which is a significant discrepancy between ability and achievement. The judge allowed the administration of intelligence tests, which measure ability, to African-American students (Salvia & Ysseldyke, 1995).

PASE v. Hannon (1980) was a class action suit brought by Parents in Action on Special Education on behalf of all African-American children who had been or would have been placed in classes for individuals with educable mental retardation in the Chicago Public Schools (CPS). The

plaintiffs noted that African-American students comprised 82% of the enrollment in classes for individuals with educable mental retardation, whereas only 62% of the total CPS enrollment was composed of African-American students (Salvia & Ysseldyke, 1995). Unlike in Larry P., the judge sided with the defendants. IQ tests were deemed to be nondiscriminatory. In addition, they were not the only bases for classification because multifaceted testing was also used (Turnbull, 1993).

Bias

The social context of assessment has changed. In the past, tests were primarily developed for use in the selection process. Today they are expected to facilitate equal opportunity. Critics have argued that tests are biased against minority groups because the unequal treatment they have received from society is not taken into consideration. Thus, the abilities of minority groups are not accurately assessed (Gordon & Terrell, 1981).

At the heart of the issue is culture. IQ tests have been criticized as biased against groups of students from deprived and culturally different backgrounds. Students with limited English language proficiency and those with disabilities are among these groups. Intelligence tests need to be culture-fair, i.e., the equity of all students

irrespective of their cultures and life experiences. The goal of the test is to limit content to material that is common to all cultures. But culture-fair tests are difficult to develop. Still, some test writers have tried to reduce "culturally loaded items" in tests, e.g., picture or vocabulary that may be biased against particular groups (Venn, 1994).

Murphy and Davidshofer (1991) established the two general purposes of tests: "to measure a particular characteristic or attribute, and to predict scores on some criterion or outcome measure" (p. 258). Based on these two uses, the major types of test bias were distinguished: "Bias in measurement occurs when the test makes systematic errors in measuring a specific characteristic or attribute . . . . Bias in prediction, on the other hand, occurs when the test makes systematic errors in predicting some criterion or outcome" (p. 258).

Using inferences from four models of nonbiased tests, LaGrow and Prochnow-LaGrow (1985) considered bias in tests of assessment procedures to exist for a subgroup under the following conditions:

> (1) Consistent test scores much lower than would be predicted from the norms of the test; (2) consistent selection for success less often than their likelihood of success on the criterion task would predict; (3) consistent selection in smaller proportions than the proportion of

> candidates who would actually be successful in meeting the criterion for success due to cultural differences; and (4) members who reach the criteria for selection are selected proportionately less often than members of other groups. (p. 34)

In reviewing statistical approaches for assessing measurement bias, Millsap and Everson (1993) suggested further understanding of construct validity, i.e., "the extent to which the test in question may be said to measure a given psychological or educational construct or trait" (Reynolds, 1991, p. 22). Bias in a given measure exists because the constructs being measured are not fully understood. Thus, studies of measurement bias should be included in the general process of construct validation (Millsap & Everson, 1993).

Bias detection has been performed on intelligence tests. Vance and Sabatino (1991) reviewed the validity bias as related to the Wechsler Intelligence Scale for Children-Revised (WISC-R) (Wechsler, 1974). They concluded that most objective data did not support bias of WISC-R. The study proposed that test items should be drawn from current cultural and linguistic information that influence child development. However, such items may be disadvantageous to children who are not exposed to this information.

Examining the use of an instrument may be advantageous than the instrument itself (Taylor, 1991). However, the use of IQ tests alone is controversial. Hilliard (1990) advocated the elimination of IQ testing in schools because it serves no purpose except to rank, classify, label, and sort based on individual's rank in a distribution of global measures of "intelligence." IQ tests are not used for diagnosis because they do not help make meaningful and valid educational decision making. They should be used only if they have demonstrated instructional validity. According to Mensh and Mensh (1991), placement of children in suitable settings by using IQ tests is based on the following assumptions: (a) IQ tests measure individual differences, and (b) assessment precedes education. As a result, using IQ tests in schools is tantamount to passing judgment on children based on racial and class lines. They explained:

> The significance of the second assumption becomes apparent when one recognizes that the tests are not instruments for assessing individual differences, but a means for ignoring individuality and slotting children according to prior assumptions about the races and classes they belong to. By using the tests to assess the children--i.e., to decide whether or not a child should be given a high-quality academic education--the schools can continue, ad infinitum, to justify superior and inferior education class and racial lines. (p. 158)

Despite accusations of bias, there are positive attitudes about IQ tests and their use. Snyderman and Rothman (1987) randomly surveyed psychologists and educational specialists (N = 1,200) from different organizations. Their findings revealed that the respondents believed that intelligence and aptitude tests adequately measure the most important elements of intelligence, which they deemed to be significant to success in society. They felt that the tests are helpful as decision-making tools in schools. Although they viewed the tests as somewhat racially and socioeconomically biased, they were apparently not influenced by political ideology.

Most norm-referenced measures, including intelligence, achievement, and screening instruments, are not validated for use on individuals with disabilities. Fuchs, Fuchs, Benowitz, and Barringer (1987) examined user manuals and technical supplements of 27 aptitude and achievement tests. Results indicated that minimal data were provided on the appropriateness of these tests on children with disabilities. Norm-referenced tests that have no validation data on people with disabilities should not be used on this group for diagnostic, classification, placement, and evaluation purposes.

Several authors (Fuchs, 1987; Grossman & Franklin, 1988; Reilly, 1991) have pointed out that bias exists in the decision-making process. O'Reilly, Nortrhcraft, and Sabers (1988) studied confirmation bias in special education eligibility decisions. School psychologists (N = 40) volunteered for the study, with each evaluating a simulated psychological report about a fifth-grade female. The reports contained assessment results and background information that were similar, except for the teacher's reason for referral. Results indicated that the school psychologists tended to refer the child to learning disabilities or gifted placement consideration based on the reason for referral. Although eligibility decisions are made by multidisciplinary teams, the perspectives of school psychologists, who should be knowledgeable in assessment, are important. Identification of strategies to make eligibility judgments objective was recommended.

Race and Intelligence

Race and intelligence have always been socially sensitive issues. IQ tests are usually used to prove or disprove opposing viewpoints.

Jensen (1969) argued that African Americans performed more poorly than European Americans on standardized IQ tests because of their genetic origin. He said that IQ was

hereditary, and compensatory education was not beneficial. Thus, it would be better to focus on skills in which African-American children were biologically adapted. Using voluminous data, Jensen (1980) ruled out that test bias was a possible cause of low IQ scores.

Dissenting opinions were expressed following the publication of the study. Jensen was accused of being a racist. Snyderman and Rothman (1988) defended Jensen's views. First, they pointed out that the genetic influence on group difference in IQ is changeable. Differences can be narrowed or even eliminated, for example, by environment. Second, the issue is one of differences in IQ, not inferiority of one group compared to another.

Herrnstein and Murray (1994) reiterated and elaborated Jensen's position that ethnic differences exist in cognitive ability. On average, the mean of African Americans on intelligence tests was 85, whereas the European Americans' mean was 100. The standard deviation was 15. Genes may explain ethnic differences rather than bias. Intelligence is an ethnic characteristic that is inherent. A cognitive elite, which is composed of people with more intelligence, has emerged to the top end of society, whereas people with less intelligence continue to remain on the lower end. Based on the proposition that cannot be improved, it has

been suggested that government programs on affirmative action, welfare, and remedial programs be eliminated. The underlying rationale of such view is that the existence of these programs only encourages the proliferation of people with less intelligence.

Environment can influence IQ. Scarr and Weinberg (1976) studied African-American and biracial children (N = 130), ranging in age from 4 to 12, who were adopted by affluent European-American families. The African-American children in the study had an average IQ of 106, which was higher than the average IQ of European-American children. In contrast, the average IQ score of African-American children in the region was 90. Also, the study found that the IQ scores were higher when the children were adopted earlier. It was concluded that social environment shapes IQ, especially when children are reared in the culture of IQ.

Racial differences in IQ cannot be simply traced to genetic or environmental origin. Such approach is limited because insights into causes are not provided. An appropriate approach is to study the mechanisms and influences involved in the question of the detailed causal explanation, which becomes part of a broader question of intellectual level in general. Even if genetic and

environmental factors are used in the explanation for racial differences, ideological divisions are prevented (Mackenzie, 1984).

The debate about whether intelligence is genetically or environmentally determined has nothing to do with group differences. At issue is whether intelligence can be changed, which is incongruous with the issue of heritability. Many inherited characteristics are changeable; conversely, there are many environmental characteristics that resist change. Educational, psychological, cultural, and economic types of interventions that are directed to the entire community can help eliminate differences in intelligence between African Americans and European Americans (Angoff, 1988).

Scarr (1988) noted that race and gender as psychological variables were seemingly an afterthought. Samples of African Americans and European Americans, males and females, may be reflective of differences rather than representatives of the population. In many studies, race and gender were psychologically analyzed under the research on children, social class, or parenting practices, which all vary by race and gender. The hesitation is pointless. Forthright studies of racial and gender differences are needed. Direct questions about the nature and origins of

racial and gender differences will provide information about underrepresented groups' strengths and ways of functioning. Accordingly, they can be helped to succeed in society.

Alternatives

The root of the IQ problem is the definition of intelligence, which is difficult to achieve. Gardner (1991) presented an alternative definition in his Theory of Multiple Intelligences. Individuals vary in the strength of the following intelligences: (a) logical-mathematical; (b) linguistic; (c) musical; (d) spatial; (e) bodily-kinesthetic; (f) interpersonal; and (g) intrapersonal. Intelligences are called and combined in distinct ways when performing various tasks and solving different problems. The present educational system gives emphasis only to linguistic and logical-quantitative modes of instruction.

Goleman (1995) emphasized emotional intelligence, which involves self-control, zeal and persistence, and self-motivation. A separate type, emotional intelligence does not oppose IQ. However, people's success can be better predicted on their emotional intelligence rather than their IQ. Emotional intelligence offers more fully human qualities, and it can be taught. For example,

self-restraint and compassion, which are important qualities in society, can be nurtured to complement anyone's intellectual potential.

Many school districts had reduced minority overrepresentation in special education by removing students above the cutoff of IQ 69. Others used alternative assessments and adjusted scores to accommodate environmental deficits (Heaston, 1987). The SFUSD tried a nonbiased approach to cognitive assessment, which had the following main features: (a) standardized psychometric tests were not used in any of the assessments; (b) the presence of a disability was documented; and (c) the process of learning rather than the product of learning was emphasized. Different individuals assessed students in several stages, e.g., home curriculum. In the last stage, there was an option to use a clinical/cognitive assessment, which stressed information processing. Intelligence was understood in terms of mental processes that are latent in observable behavior. Quantitative units or scores, which could be translatable into a classification, were not derived from the assessment process. Instead, the professional judgment of the assessor determined the eligibility of a student for a special education placement. Less than one-third of the African-American students in

special education were involved in the nonbiased approach. African-American enrollment in special education was eventually reduced to 7.1% (Dent, 1987).

Technical inadequacy has characterized alternative assessments. Curriculum-based measurement was presented by Deno, Marston, and Tindal (1985-1986) as an alternative to educational decision making. Students are directly and frequently measured based on their performance in the school curriculum. Curriculum-based measurement can be used in special education for screening, eligibility, program planning, progress monitoring, and program evaluation. Shinn (1988) focused on the appropriateness of the curriculum-based measurement on students with mild disabilities, i.e., learning disabilities, mild mental retardation, and emotional disturbance. Although the development of local norms are provided, curriculum-based measurement may engender labelling, misinterpretation, and mediocrity.

Dynamic assessment includes a learning part in the testing situation. The learner's responsiveness to teaching is analyzed. Several models of dynamic assessment exist. For example, in the the test-train-test assessment, which is the basic model, a child is pretested, trained, and posttested. Training levels off differences in children's

backgrounds, benefitting those who come from disadvantaged circumstances. Unlike a static IQ score, the posttest score reflects the child's opportunity to achieve success. A major drawback of dynamic assessment is construct inconsistency--the models differ in definition, theoretical foundation, and procedural requirements (Jitendra & Kameenui, 1993).

The Future

Prevention of future litigation is important. Court cases involve expensive legal fees, which are reallocated from educational budgets. The implicit assumptions and issues of past litigation have to be addressed. What has worked in the past, e.g., disproving allegations of IQ-test biases, cannot be simply adopted as a matter of course. In special education, assessments should be directed towards effective interventions rather than simple classification of students (Reschly, 1991). Assessment activities should result in enhanced instructional opportunities for students (Ysseldyke et al., 1992).

Classification practices raise constitutional questions. First, there may be a denial of equal protection when students are excluded and assigned to special education programs. Second, modifications have to made when overrepresentation of minority students in classes for

individuals with mild disabilities seemingly has racially related harmful effects. Third, misclassification can be avoided by ensuring the rights of students and their parents to a due process on placement recommendations (Kirp, 1992).

The classification system affects assessment practices. IQ testing is prevalent because the current classification system focuses on general intellectual functioning. Reschly (1988) proposed a classification system based on severe chronic achievement problems in the majority of academic subjects. Curriculum-based measures are appropriate because they can demonstrate a student's present functioning within a particular educational setting and curriculum.

Wood (1992) stated that parental involvement and improved nondiscriminatory test practices may prevent the overrepresentation of minority students in special education. The numbers of misdiagnosed students may also be decreased by prereferral intervention strategies. On all of these, what is essential is the perspective that students are placed in special education to prepare them for eventual return to general education. As such, the exit criteria from special education have to be developed and implemented well.

Reschly (1991) projected that the use of IQ tests will decline. Bias, however, is not the reason. IQ tests cannot

be used to develop interventions for individuals with learning and behavioral difficulties. Functional assessments, e.g., curriculum-based measures, will be used more because they are relevant to the design, implementation, and evaluation of interventions.

The relevance of assessment to intervention and instruction becomes critical in mainstreaming, which requires modification of instruction and evaluation of student progress. Another movement is inclusion, i.e., educating students with disabilities in the general education setting (Salend, 1994). While taking into consideration their differences, the needs of all students are served into a single educational system that unites general education and special education (Stainback & Stainback, 1986). The focus is on effective instruction (Gartner & Lipsky, 1987), which is adapted for students with mild disabilities in general education classes (Ysseldyke et al., 1992).

References

Angoff, W. H. (1988). The nature-nature debate, aptitudes, and group differences. American Psychologist, 43(9), 713-720.

Beirne-Smith, M., Patton, J., & Ittenbach, R. (1994). Mental retardation (4th ed.). New York: Merrill.

Deno, S. L., Marston, D., & Tindal, G. (1985-1986). Direct and frequent curriculum-based measurement: An alternative for educational decision making. Special Services in the Schools, 2(2-3), 5-27.

Dent, H. E. (1987). The San Francisco public schools experience with alternatives to I.Q. testing: A model for non-biased assessment. Negro Educational Review, 38(2-3), 146-162.

Diana v. State Board of Education, Civ. Act No. C-70-37 (N.D. Cal. 1970).

Fuchs, D. (1987). Examiner familiarity effects on test performance: Implications for training and practice. Topics in Early Childhood Special Education, 7(3), 90-104.

Fuchs, D. Fuchs, L. S., Benowitz, S., & Barringer, K. (1987). Norm-referenced tests: Are they valid for use with handicapped students? Exceptional Children, 54(3), 263-271.

Gardner, H. (1991). The unschooled mind: How children think and how schools teach. New York: Basic Books.

Gartner, A., & Lipsky, D. K. (1987). Beyond special education: Toward a quality system for all students. Harvard Educational Review, 57(4), 367-395.

Goleman, D. (1995). Emotional intelligence. New York: Bantam Books.

Gordon, E. W., & Terrell, M. D. (1981). American Psychologist, 36(10), 1167-1171.

Grossman, H. J. (1983). Classification in mental retardation. Washington, DC: American Association on Mental Deficiency.

Grossman, F. M., & Franklin, N. K. (1988). Bias effects in speech-language assessment and decision-making. Language, Speech, and Hearing Services in Schools, 19(2), 153-159.

Heaston, P. (1987). The Chicago Public Schools experience with alternatives to I.Q. testing: The color of rubies. Negro Educational Review, 38(2-3), 163-172.

Herrnstein, R. J., & Murray, C. (1994). The bell curve: Intelligence and class structure in American life. New York: Free Press.

Hilliard, A. G., III. (1990). The case against the use of IQ tests in the schools. Contemporary Education, 61(4), 184-189.

Hunt, N., & Marshall, K. (1994). Exceptional children and youth: An introduction to special education. Boston: Houghton Mifflin.

Incidence of racial isolation in California schools: Office of Intergroup Relations report. (1983). Sacramento, CA: California State Department of Education.

Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? Harvard Educational Review, 39(1), 1-123.

Jensen, A. R. (1980). Bias in mental testing. New York: Free Press.

Jitendra, A. K., & Kameenui, E. J. (1993). Dynamic assessment as a compensatory assessment approach: A description and analysis. Remedial and Special Education, 14(6), 6-18.

Kirp, D. L. (1992). Student classification, public policy, and the courts. In T. Hehir & T. Latus (Eds.), Special education at the century's end: Evolution of theory and practice since 1970 (pp. 3-45). Cambridge, MA: Harvard Educational Review.

LaGrow, S. J., & Prochnow-LaGrow, J. E. (1985). Consideration of bias in the assessment and placement process of exceptional children. In A. F. Rotatori & R. Fox

(Eds.), Assessment for regular and special education for teachers: A case study approach (pp. 31-51). Austin, TX: PRO-ED.

Larry P. v. Riles, C-71-2270-RFP (N.D. Cal. 1972), 495 F. Supp. 96 (N.D. Cal. 1979), aff'd, (9th Cir. 1984), 1983-84 EHLR DEC. 555: 304.

Lora v. Board of Education of the City of New York, 74 F.R.D. 565 (E.D. N.Y. 1977); 456 F. Supp. 1211 (E.D. N.Y 1978), remanded, 623 F.2d 248 (2nd Cir. 1980); 587 F. Supp. 1572 (E.D. N.Y. 1984).

Mackenzie, B. (1984). Explaining race differences in IQ: The logic, the methodology, and the evidence. American Psychologist, 39(11), 1214-1233.

MacMillan, D. L., & Balow, I. H. (1991). Impact of Larry P. on educational programs and assessment practices in California. Diagnostique, 17(1), 57-69.

MacMillan, D. L., Hendrick, I. G., & Watkins, A, V. (1988). Impact of Diana, Larry P., and P.L. 94-142 on minority students. Exceptional Children, 54(5), 426-432

Maheady, L., Towne, R., Algozzine, B., Mercer, J., & Ysseldyke, J. (1990). Minority overrepresentation: A case for alternative practices prior to referral. In S. B. Sigmon (Ed.), Critical voices on special education: Problems and

progress concerning the mildly handicapped (pp. 89-102). Albany, NY: State University of New York Press.

McLoughlin, J. A., & Lewis, R. B. (1994). Assessing special students (4th ed.). New York: Merrill.

Mensh, E., & Mensh, H. (1991). The IQ mythology: Class, race, gender, and inequality. Carbondale, IL: Southern Illinois University Press.

Mercer, C. D. (1991). Students with learning disabilities (4th ed.). New York: Merrill.

Millsap, R. E., & Everson, H. T. (1993). Methodology review: Statistical approaches for assessing measurement bias. Applied Psychological Measurement, 17(4), 297-334.

Murphy, K. R., & Davidshofer, C. O. (1991). Psychological testing: Principles and applications (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

O'Reilly, C., Northcraft, G. B., & Sabers, D. (1988). The confirmation bias in special education eligibility decisions. School Psychology Review, 18(1), 126-135.

Overton, T. (1992). Assessment in special education: An applied approach. New York: Merrill.

PASE (Parents in Action in Special Education) v. Hannon, 506 F. Supp. 831 (N.D. Ill. 1980).

Podemski, R. S., Marsh, G. E., II, Smith, T. E. C., & Price, B. J. (1995). Comprehensive administration of special education (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Polloway, E. A., & Smith, J. D. (1988). Current status of the mild mental retardation construct: Identification, placement, and programs. In M. C. Wang, M. C. Reynolds, & H. J. Walberg (Eds.), Handbook of special education: Research and practice: Vol. 2. Mildly handicapped conditions (pp. 7-22). Oxford, England: Pergamon.

Prasse, D. P. (1988). Legal influence and educational policy in special education. Exceptional Children, 54(4), 302-308.

Prasse, D. P., & Reschly, D. J. (1986). Larry P.: A case of segregation, testing, or program efficacy? Exceptional Children, 52(4), 333-346.

Reilly, T. F. (1991). Cultural bias: The albatross of assessing behavior-disordered children and youth. Preventing School Failure, 36(1), 50-53.

Reschly, D. J. (1988). Assessment issues, placement litigation, and the future of mild mental retardation classification and programming. Education and Training in Mental Retardation, 23(4), 285-301.

Reschly, D. J. (1991). Bias in cognitive assessment: Implications for future litigation and professional practices. Diagnostique, 17(1), 86-90.

Reschly, D. J., & Grimes, J. P. (1990). Best practices in intellectual assessment. In A. Thomas & J. Grimes (Eds.), Best practices in school psychology-II (pp. 425-439). Washington, DC: National Association of School Psychologists.

Reschly, D. J., Kicklighter, R., & McKee, P. (1988). Recent placement litigation Part III: Analysis of differences in Larry P., Marshall, and S-1 and implications for future practices. School Psychology Review, 17(1), 39-50.

Reynolds, C. R. (1991). Methods for studying bias in psychological and educational tests. Diagnostique, 17(1), 21-39.

Salend, S. J. (1994). Effective mainstreaming: Creating inclusive classrooms (2nd ed.). New York: Macmillan.

Salvia, J., & Ysseldyke, J. E. (1995). Assessment (6th ed.). Boston: Houghton Mifflin.

Scarr, S. (1988). Race and gender as psychological variables: Social and ethical issues. American Psychologist, 43(1), 56-59.

Scarr, S., & Weinberg, R. A. (1976). IQ test performance of black children adopted by white families. American Psychologist, 31(10), 726-739.

Shinn, M. R. (1988). Development of curriculum-based local norms for use in special education decision-making. School Psychology Review, 17(1), 61-80.

Siegel, L. S. (1989). IQ is irrelevant to the definition of learning disabilities. Journal of Learning Disabilities, 22(8), 469-478, 486.

Snyderman, M., & Rothman, S. (1987). Survey of expert opinion on intelligence and aptitude testing. American Psychologist, 42(2), 137-144.

Snyderman, M., & Rothman, S. (1988). The IQ controversy, the media and public policy. New Brunswick, NJ: Transaction.

Stainback, W., & Stainback, S. (1986). One system, one purpose: The integration of special and regular education. Entourage, 1(3), 12-16.

Taylor, R. L. (1991). Bias in cognitive assessment: Issues, implications, and future directions. Diagnostique, 17(1), 3-5.

Turnbull, H. R., III. (1993). Free appropriate public education: The law and children with disabilities (4th ed.). Denver, CO: Love.

Underwood, J. K., & Mead, J. F. (1995). Legal aspects of special education and pupil services. Boston: Allyn and Bacon.

U.S. Department of Commerce. (1994). Statistical abstract of the United States 1994 (114th ed.). Washington, DC: Author.

Vance, B., & Sabatino, D. (1991). Identifying sources of bias in the WISC-R. Diagnostique, 17(1), 40-48.

Venn, J. (1994). Assessment of students with special needs. New York: Merrill.

Weschler, D. (1974). Weschler Intelligence Scale for Children-Revised. San Antonio, TX: Psychological Corporation.

Wood, J. W. (1992). Adapting instruction for mainstreamed and at-risk students (2nd ed.). New York: Merrill.

Wood, F. H., Johnson, J. L., & Jenkins, J. R. (1986). The Lora case: Nonbiased referral, assessment, and placement procedures. Exceptional Children, 52(4), 323-331.

Ysseldyke, J. E., Algozzine, B., & Thurlow, M. L. (1992). Critical issues in special education (2nd ed.). Boston, MA: Houghton Mifflin.